Beyond generative models: superfast traversal, optimization, novelty, exploration and discovery (STONED) algorithm for molecules using SELFIES

Year: 2020 | Citations: 134 | Authors: AkshatKumar Nigam, R. Pollice, Mario Krenn, Gabriel dos Passos Gomes, Alán Aspuru-G

Abstract

Inverse design allows the generation of molecules with desirable physical quantities using property optimization. Deep generative models have recently been applied to tackle inverse design, as they possess the ability to optimize molecular properties directly through structure modification using gradients. While the ability to carry out direct property optimizations is promising, the use of generative deep learning models to solve practical problems requires large amounts of data and is very time-consuming. In this work, we propose STONED – a simple and efficient algorithm to perform interpolation and exploration in the chemical space, comparable to deep generative models. STONED bypasses the need for large amounts of data and training times by using string modifications in the SELFIES molecular representation. First, we achieve non-trivial performance on typical benchmarks for generative models without any training. Additionally, we demonstrate applications in high-throughput virtual screening for the design of drugs, photovoltaics, and the construction of chemical paths, allowing for both property and structure-based interpolation in the chemical space. Overall, we anticipate our results to be a stepping stone for developing more sophisticated inverse design models and benchmarking tools, ultimately helping generative models achieve wider adoption.